YAMAHA
Make Waves

From classrooms to cathedrals

Enhancing Speech Intelligibility

Wouter Verkuijl
Yamaha Music Europe

Ruben van der Goor
Yamaha Music Europe

SDC 2025
Vienna

O ou i i ie o ea ?

Eaue o e eoea oe o e ua i, o eoe o.

D y fnd ths smple t rd ?

Bcs f th phnmnl pwr f th hmn mnd, mst ppl d.

YAMAHA

Do yuo fnid tihs smilpe to raed?

Bceuase of the phaonmneal pweor of the hmuan mnid, msot plepoe do.

YAMAHA

Do you find this simple to read?

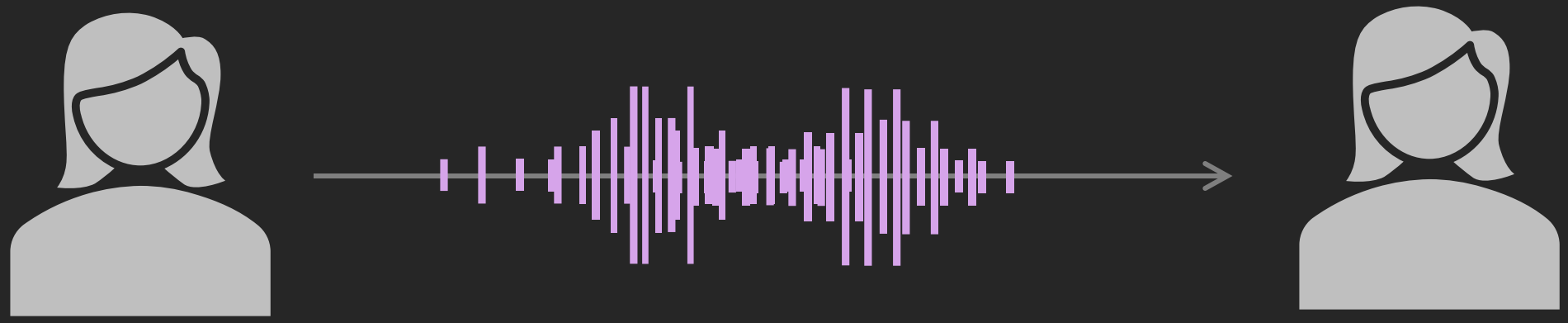
Because of the phenomenal power of the human mind, most people do.

YAMAHA

Introduction

- Why is speech intelligibility of importance?
  - Inform/communicate
  - Educate
  - Entertain
  - Decision making
  - Evacuate

How to measure “real” intelligibility ?



“First you hear ‘duk’, then I say ‘zwat’, followed by ‘keeg’”

Perception	Word	Result
	duk	☑
	zwak	☒
	keeg	☑

YAMAHA

How to measure “real” intelligibility ?

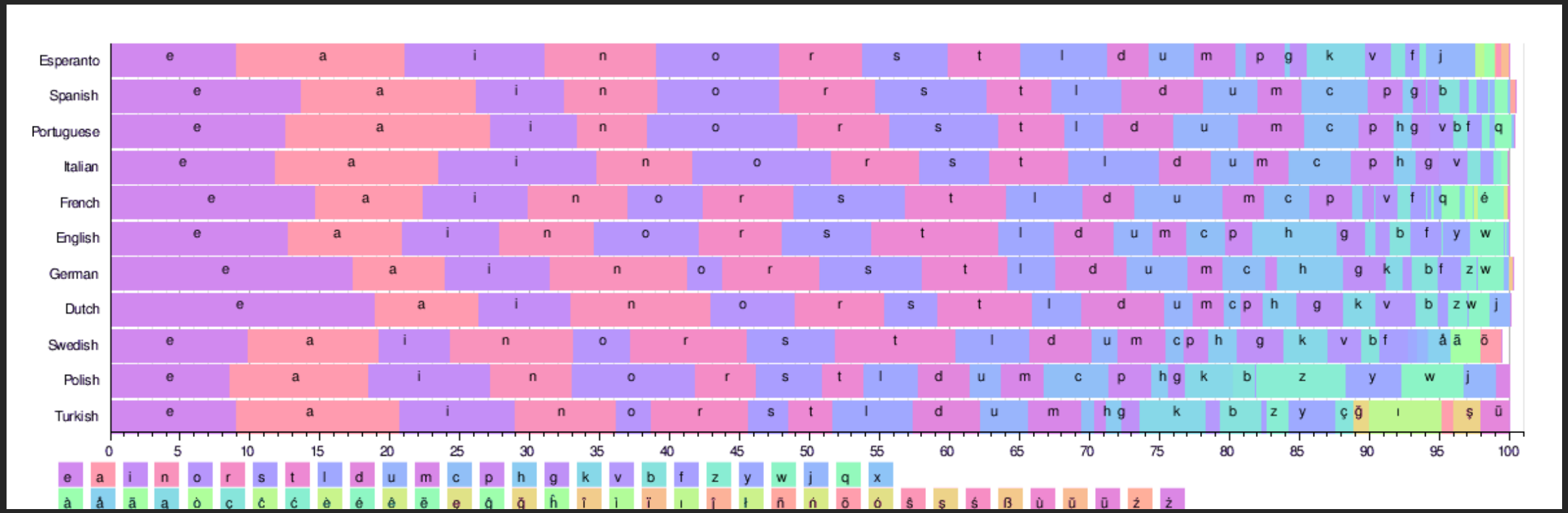
- %Alcons: Articulation Loss of Consonants (%)

	Articulation Loss of Consonants (%)
Excellent	0 - 3
Good	3 - 7
Fair	7 - 15
Poor	15 - 33
Bad	33 - 100

How to measure “real” intelligibility ?

- Consonants are more important in speech than vowels *(Cnsnnts <> ooa)*
- Consonants carry more information (depending on language)

How to measure “real” intelligibility ?



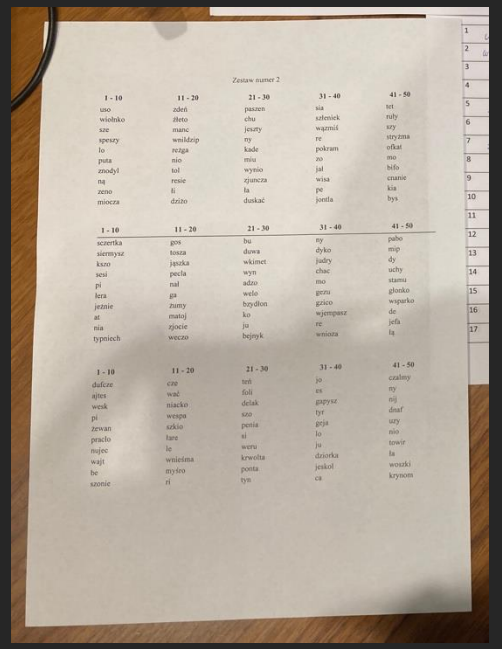
How to measure “real” intelligibility ?

- Consonants are more important in speech than vowels *(Cnsnnts <> ooa)*
- Consonants carry more information (depending on language)
- Consonants carry more energy than vowels (5-10dB)
- Most energy of consonants are in the higher frequency domain
  - Speech: 125Hz – 8kHz
  - Most important : 300Hz - 4kHz
- 15dB on top of BGN is needed for good intelligibility

How to measure “real” intelligibility ?

- Panels of live speakers and listeners
- Subjective test of reproduction of predefined syllables, words or sentences.
- Logatomes are used
- Law of large numbers: lots of effort and very time consuming.
- Practical problems
  - System availability
  - Reliable testing environment
  - Dangerous environments

How to measure “real” intelligibility ?

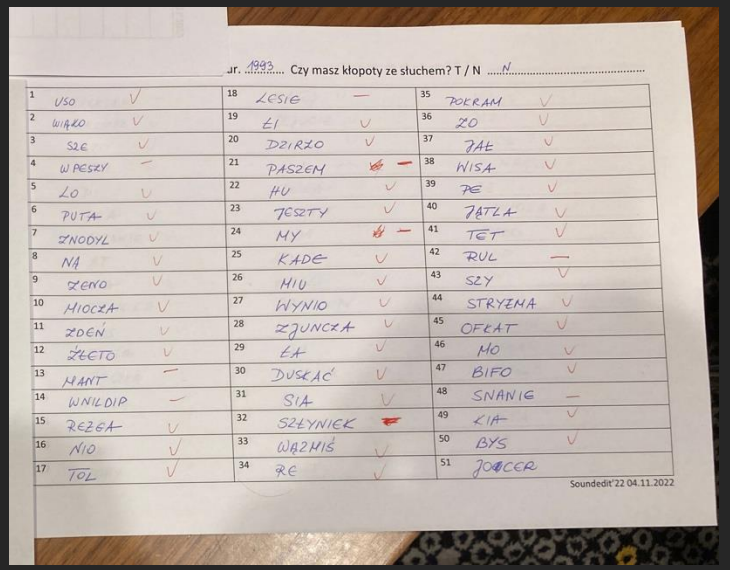


Zestaw numer 2

1 - 10	11 - 20	21 - 30	31 - 40	41 - 50
uso	zdeń	paszen	sia	tet
wiośnko	złeto	chu	szłemiek	ruły
sze	manc	jeszty	wązmić	szy
speszy	wnildzip	ny	re	stryzna
lo	rezga	kade	pokram	ofkat
puta	nio	miu	zo	mo
znodył	tol	wynio	jał	bifo
ną	resie	zjuncza	wisa	cnanie
zeno	fi	fa	pe	kia
miocza	dzizo	dukać	jonfa	bys

1 - 10	11 - 20	21 - 30	31 - 40	41 - 50
sczertka	gos	bu	ny	pabo
siernysz	tosza	duwa	dyko	mip
kszo	jązka	wkimet	judry	dy
sesi	pecla	wyn	chac	uchy
pi	nal	adzo	mo	stamu
fera	ga	welo	gezu	głonko
jeznie	zumy	bzydlon	gzico	wsparko
at	matoj	ko	wjempasz	de
nia	zjonie	ju	re	jefa
typniech	weczo	bejnyk	wnioza	łą

1 - 10	11 - 20	21 - 30	31 - 40	41 - 50
dufcze	cza	teń	jo	czalmy
ajtes	wać	foli	es	ny
wesk	niacko	delak	gapysz	nij
pi	wespa	szo	tyr	draf
zewan	szkio	penta	grzá	uzy
pracło	bare	si	lo	nio
nojec	le	weru	ju	towir
wajt	wnieśma	krwołta	dziorka	fa
be	myśro	ponta	jeskol	woszki
szonie	ri	tyn	ca	krysnom



ur. 1993 Czy masz kłopoty ze słuchem? T / N N

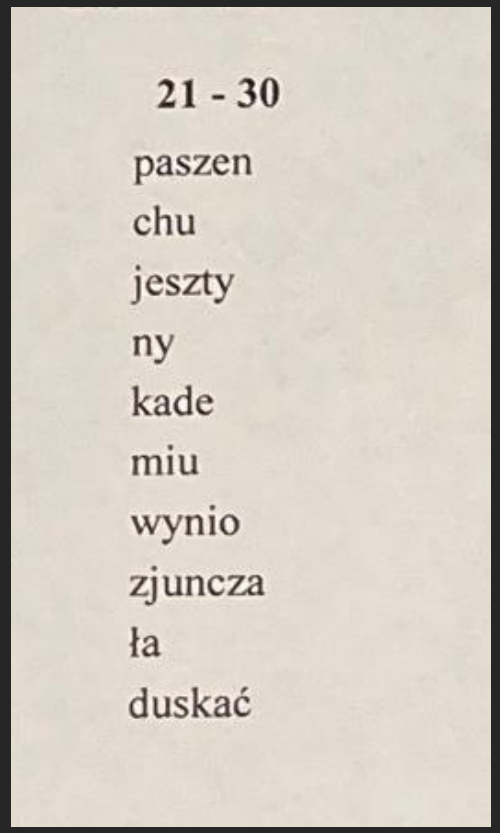
1	USO	√	18	ZESIE	—
2	WIĄKO	√	19	ŁI	√
3	SZE	√	20	DZIRZO	√
4	WPESZY	—	21	PASZEM	—
5	LO	√	22	HU	√
6	PUTA	√	23	JESZTY	√
7	ZNODYL	√	24	MY	—
8	NĄ	√	25	KADE	√
9	ZENO	√	26	MIU	√
10	MIOCZA	√	27	WYNIO	√
11	ZDEŃ	√	28	ZJUNCZA	√
12	ZŁETO	√	29	ŁA	√
13	MANT	—	30	DUSŁAĆ	√
14	WNILDIP	—	31	SIA	√
15	REZGA	√	32	SZŁYNIEK	—
16	NIO	√	33	WĄZMIŚ	√
17	TOL	√	34	RE	√
35	POKRAM	√			
36	ZO	√			
37	JAŁ	√			
38	WISA	√			
39	PE	√			
40	JĄTLA	√			
41	TĘT	√			
42	RUL	—			
43	SZY	√			
44	STRYZNA	√			
45	OFŁAT	√			
46	MO	√			
47	BIFO	√			
48	SNANIE	—			
49	ŁIA	√			
50	BYS	√			
51	JOŁCER				

Soundedit'22 04.11.2022

STI in enhanced audio systems, by Dr Pastusiak and Błasiński, SoundEdit Łodz, 2022

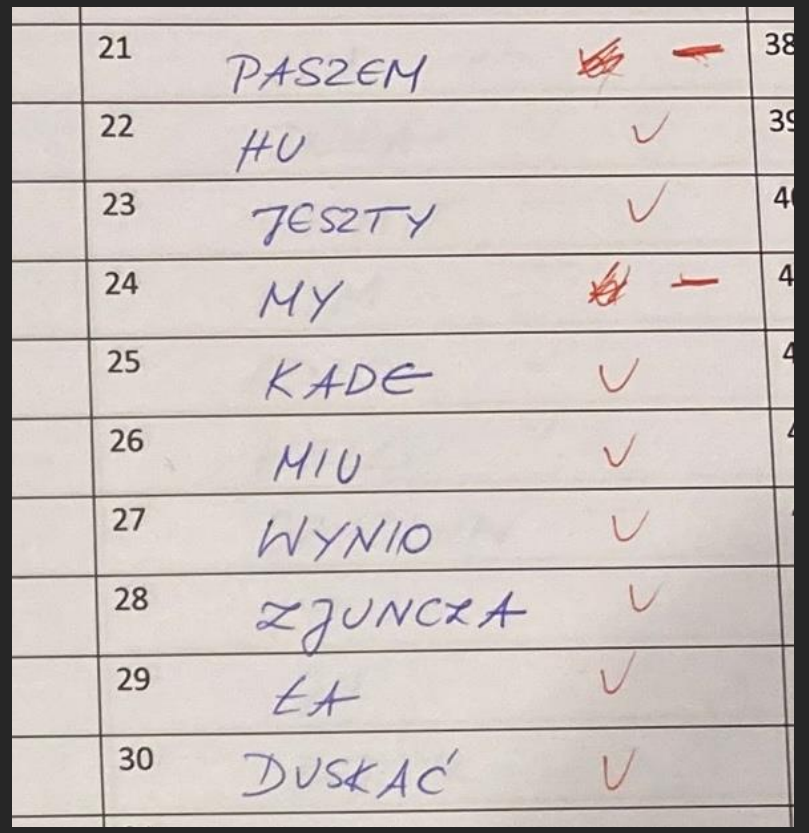
YAMAHA

How to measure “real” intelligibility ?



21 - 30

paszen
chu
jeszty
ny
kade
miu
wynio
zjuncza
ła
duskać



21	PASZEM	~~✓~~ –	38
22	HU	✓	39
23	JESZTY	✓	4(
24	MY	~~✓~~ –	4
25	KADE	✓	4
26	MIU	✓	
27	WYNIO	✓	
28	ZJUNCZA	✓	
29	ŁA	✓	
30	DUSKAĆ	✓	

STI in enhanced audio systems, by Dr Pastusiak and Błasiński, SoundEdit Łodz, 2022

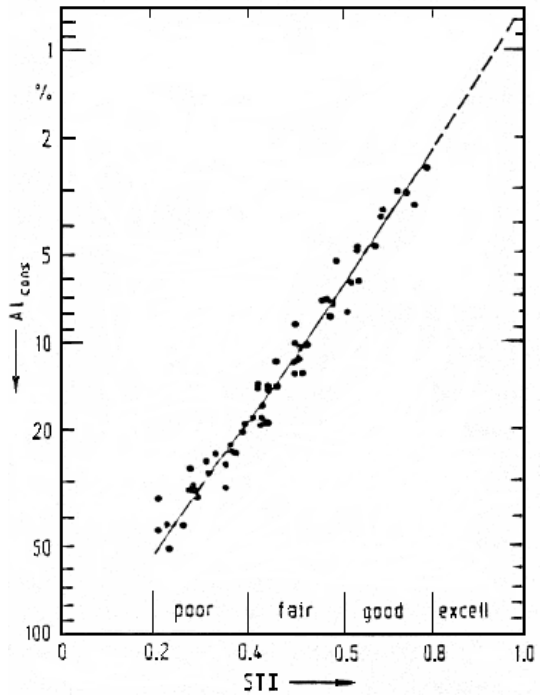
YAMAHA

What is STI ?

- STI = Speech Transmission Index
- An objective measurement tool to judge quality of speech transmission in any system (the transmission channel)
- Developed by Houtgast and Steeneken (TNO, NL) in the 70’s
- STI only describes the transmission channel, not talker/listener specifications
- Standardization
  - Standardized reference signal
  - Standardized measuring method
  - Standardized evaluation method
- All described in EN-IEC-60268-16 rev 5 (2020)
- The result is a value between 0 and 1

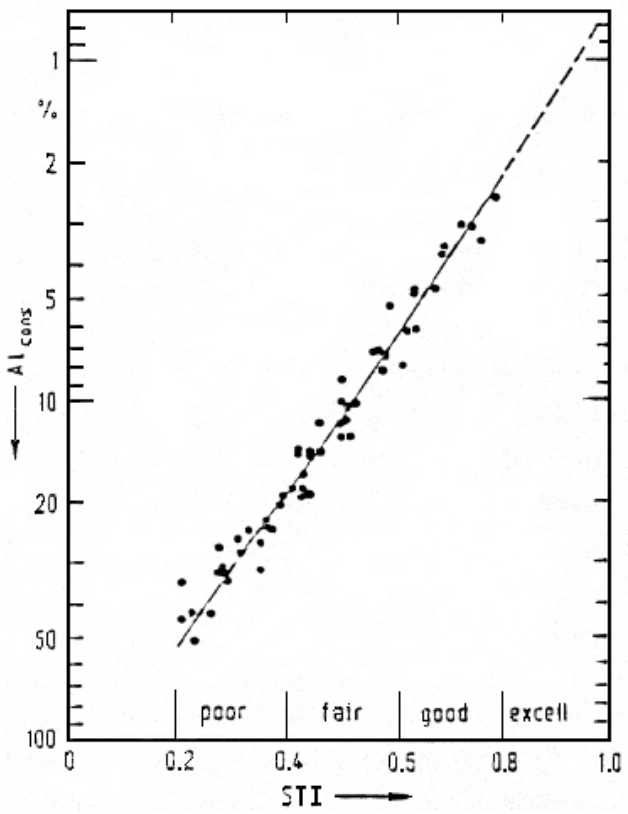
How to measure intelligibility ?

- Any means of electronically determining the intelligibility is based on statistical relation between a measured parameter and the “real” intelligibilty
- The interpreter of the acquired data must therefore always be aware of conditions and limitations.



How to measure intelligibility ?

- Relation to %ALcons



$$\%ALcons = 10^{\frac{1-STI}{0,46}}$$

$$\%ALcons = 170,5405 \cdot e^{-5,419 \cdot STI}$$

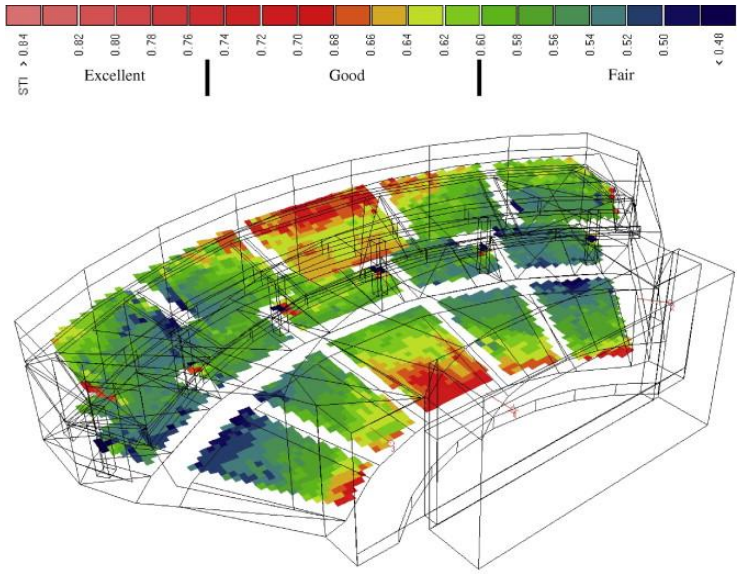
$$STI = 1 - 0,46 \cdot \log(\%ALcons)$$

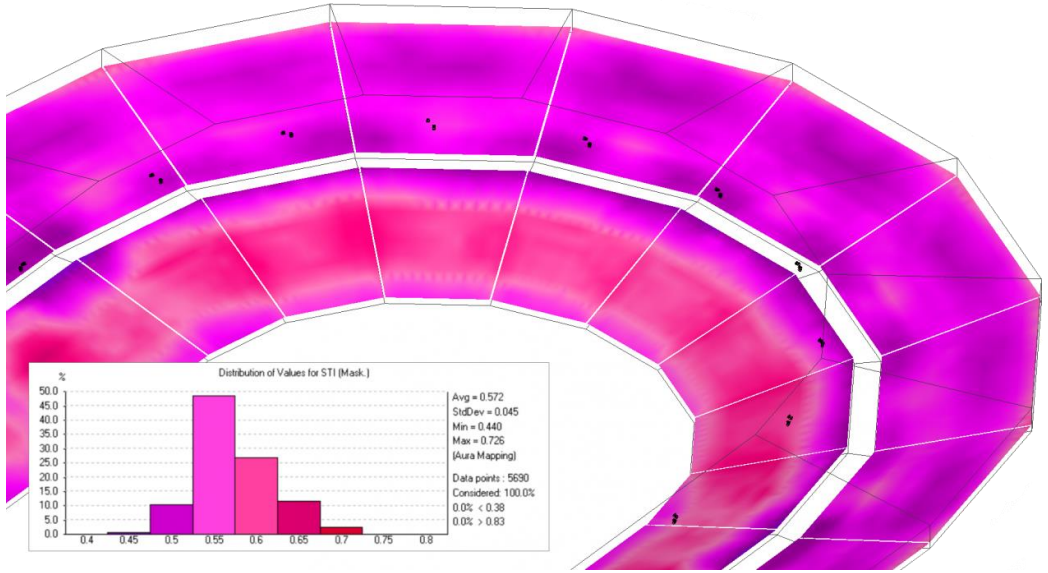
$$STI = 0,185 \cdot \ln\left(\frac{\%ALcons}{170,5405}\right)$$

YAMAHA

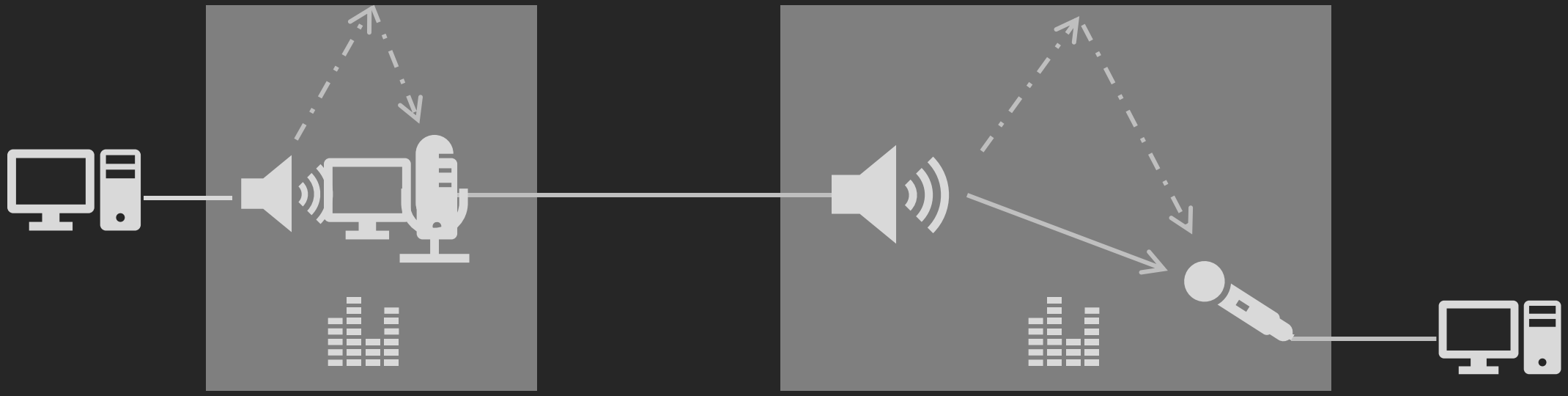
What is STI ?

- STI can be measured (Bedrock, NTI, Smaart)
- STI can be predicted (EASE, Odeon)





How to measure intelligibility ?

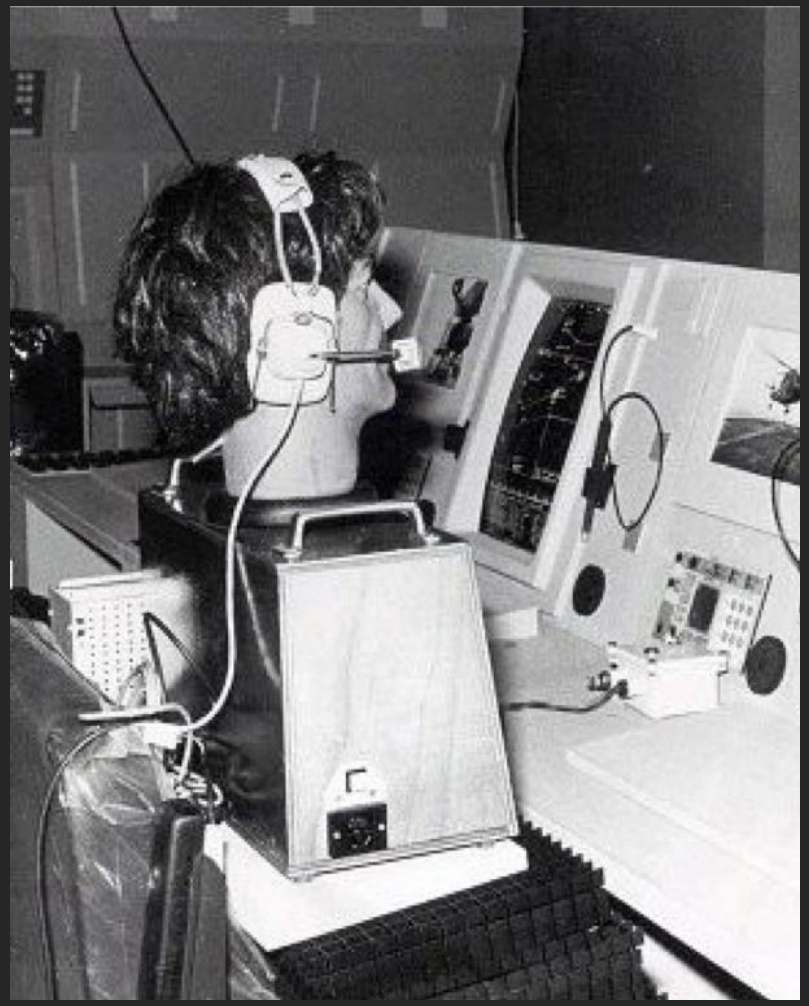


How to measure intelligibility ?

- Measuring in real life

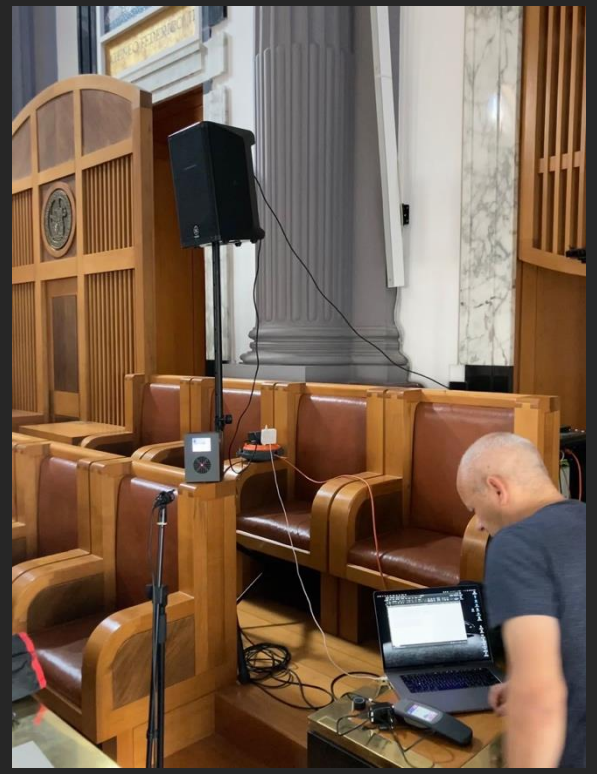
How to measure intelligibility ?

- Measuring in real life



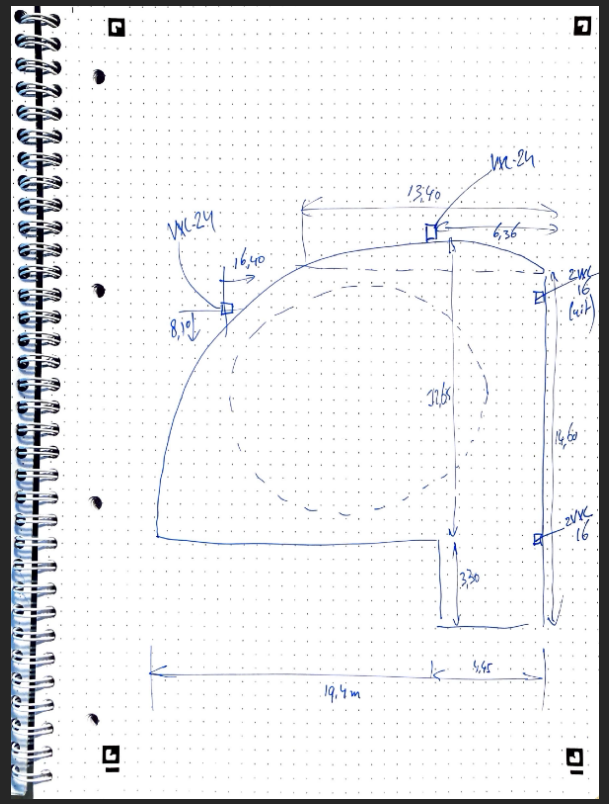
How to measure intelligibility ?

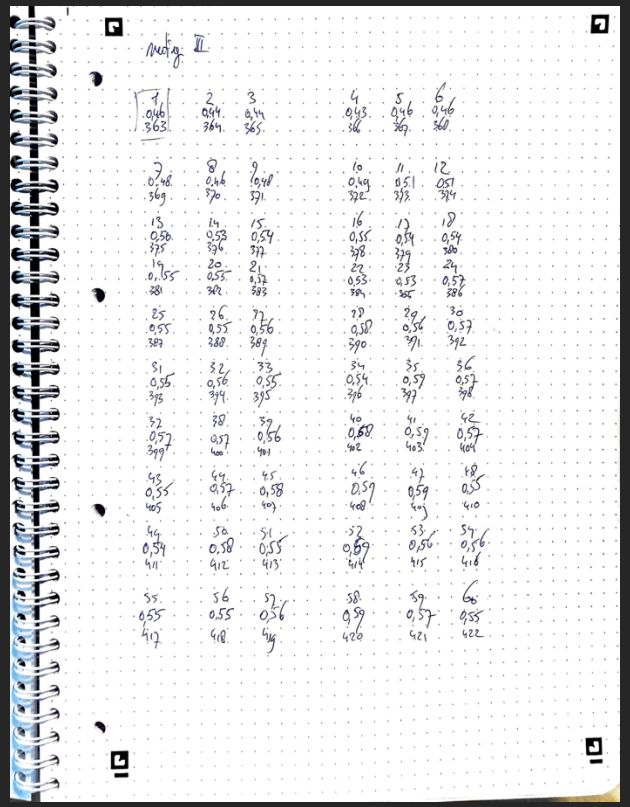
- Measuring in real life

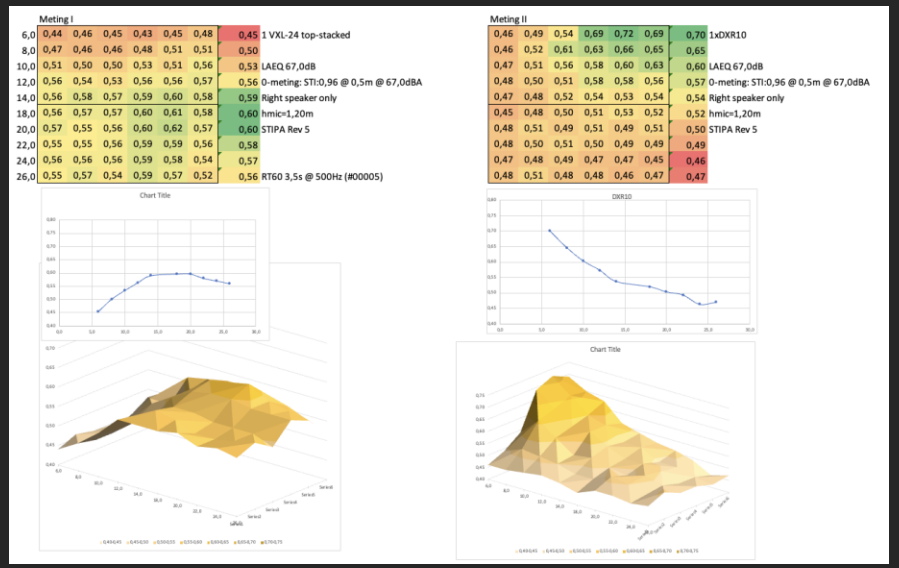


How to measure intelligibility ?

- Measuring in real life

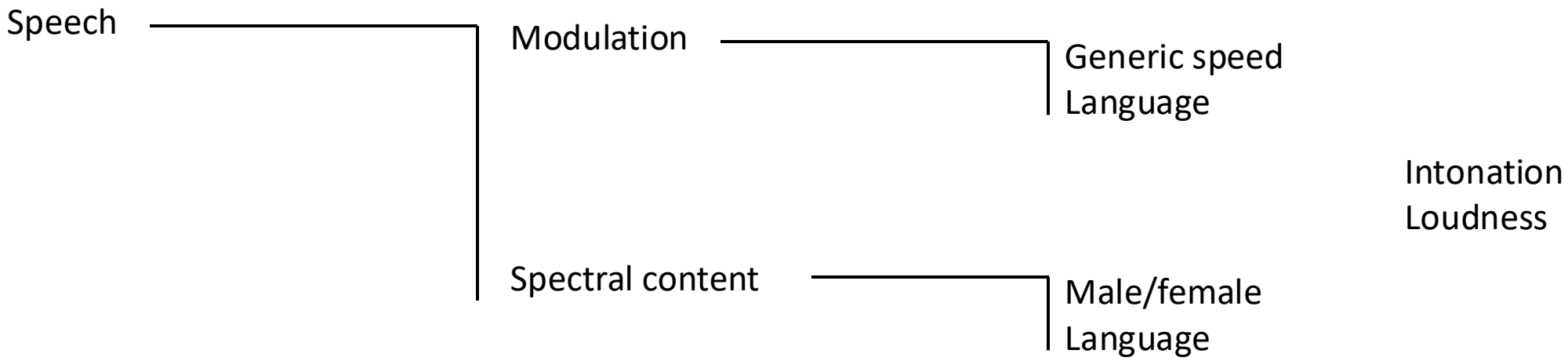






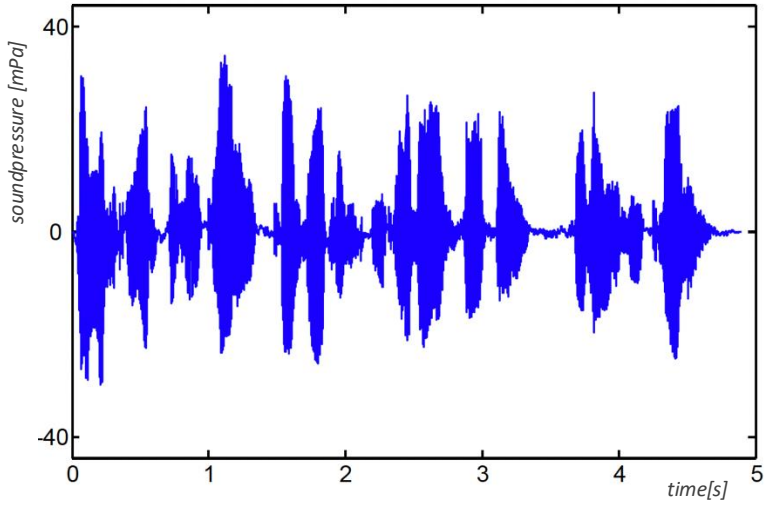
How to measure intelligibility ?

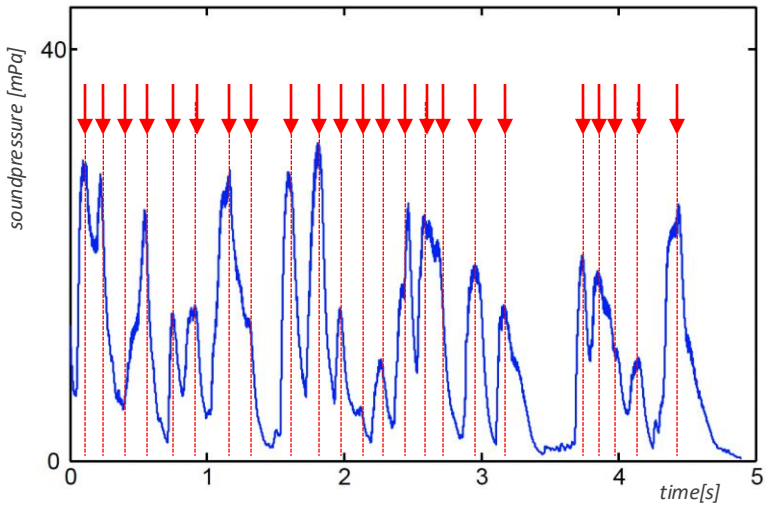
- Houtgast and Steeneken



How to measure intelligibility ?

- Modulation



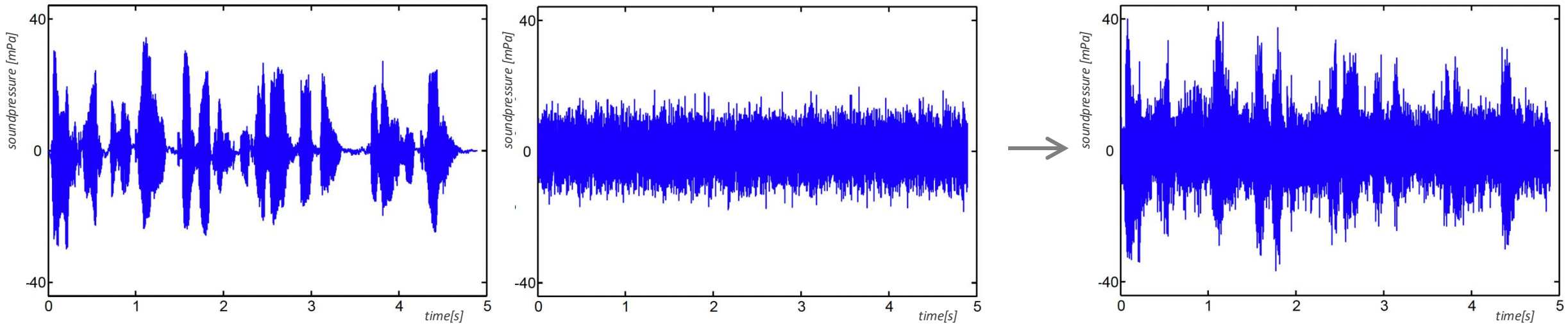


Graphics TU Delft

YAMAHA

How to measure intelligibility ?

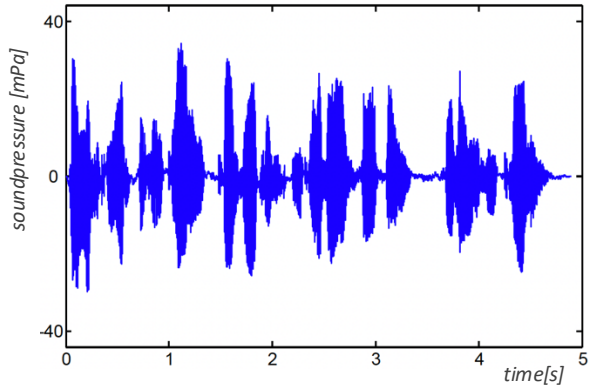
- Modulation

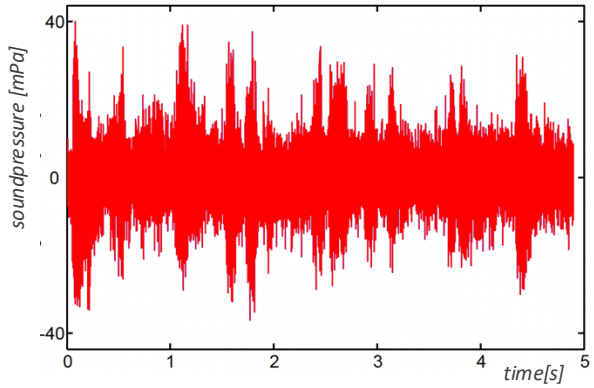


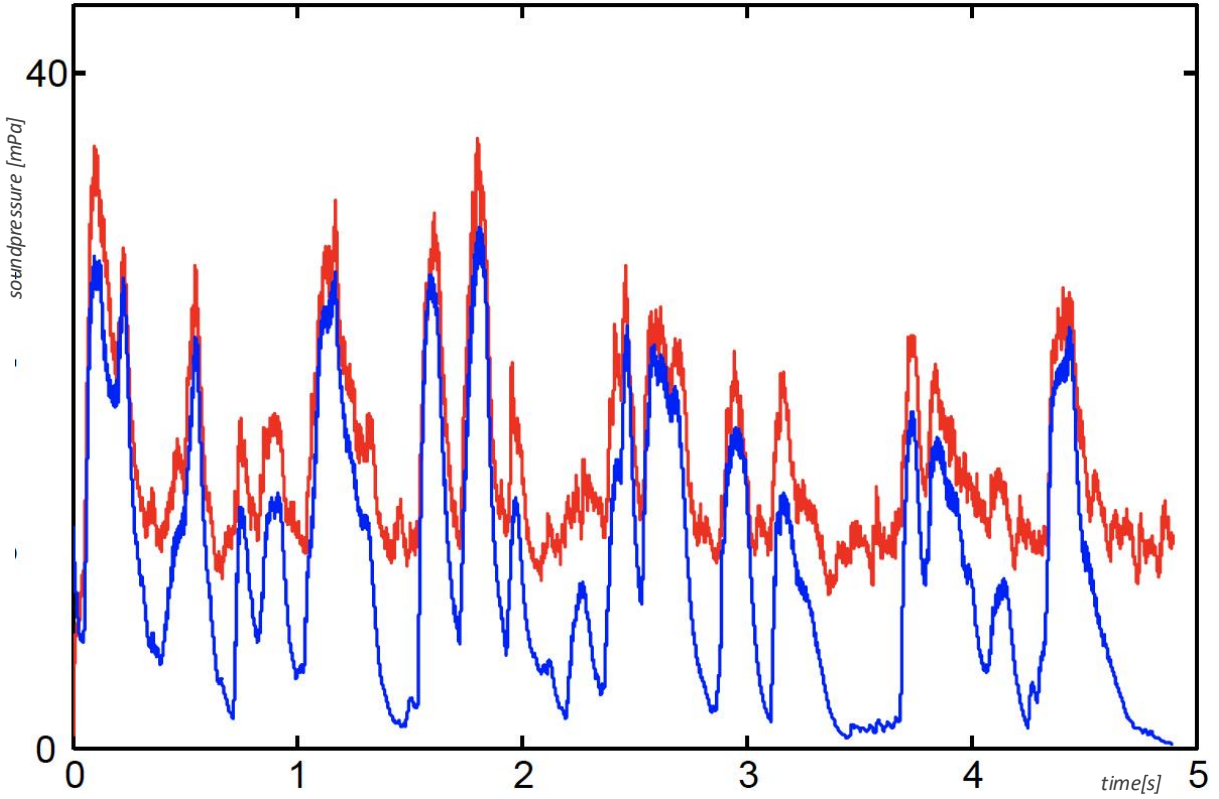
Graphics TU Delft

How to measure intelligibility ?

- Modulation



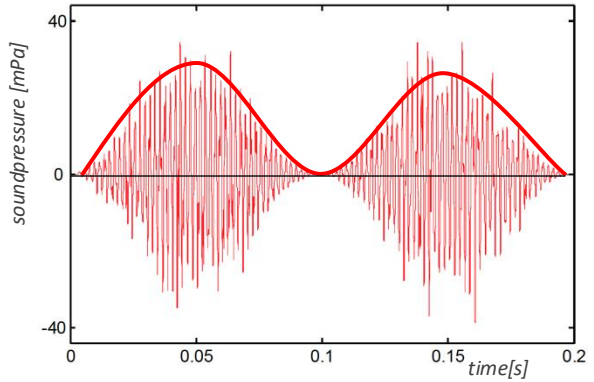


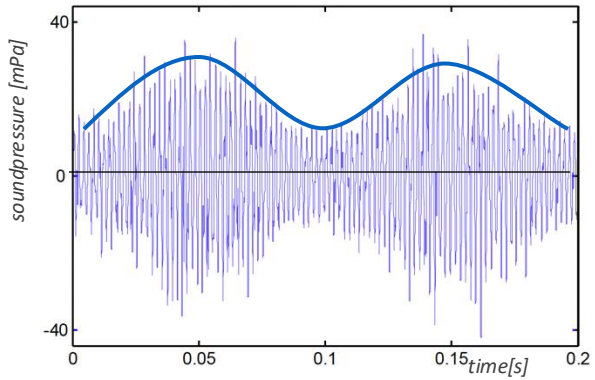


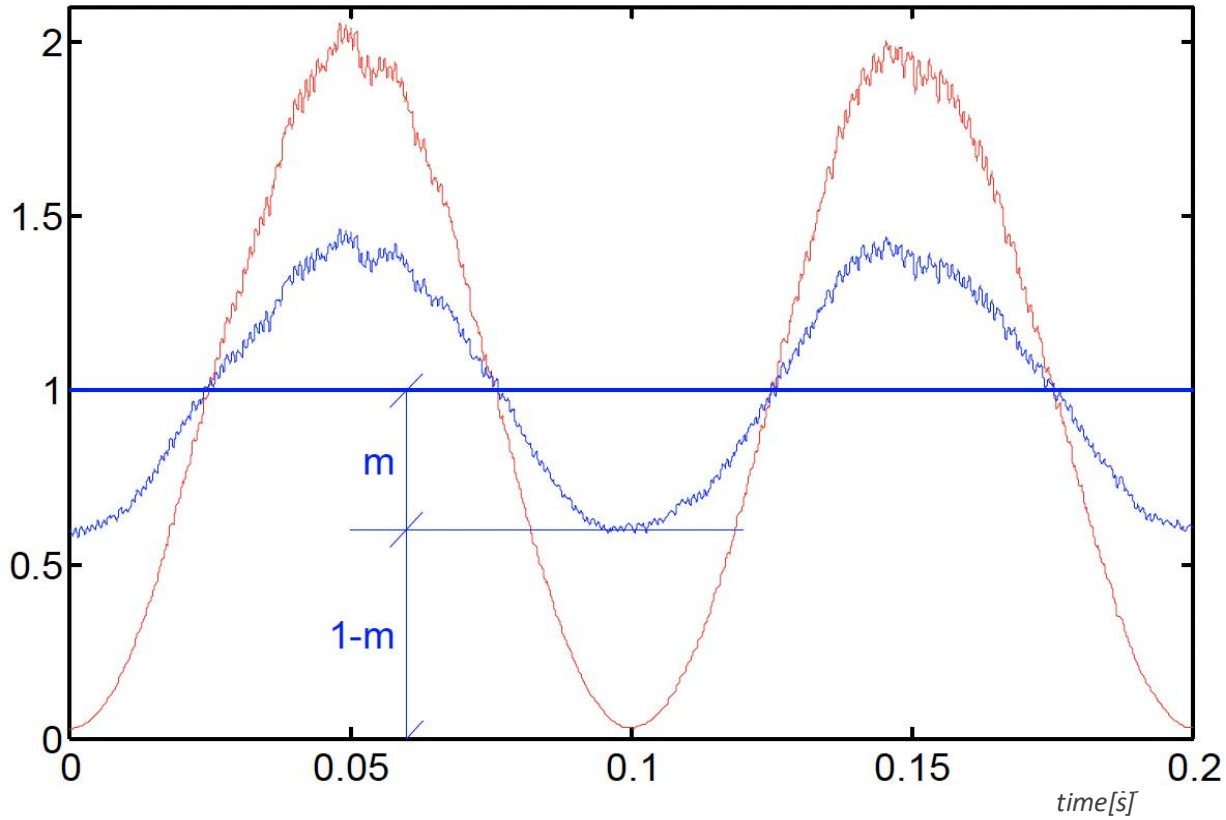
Graphics TU Delft

How to measure intelligibility ?

- Modulation



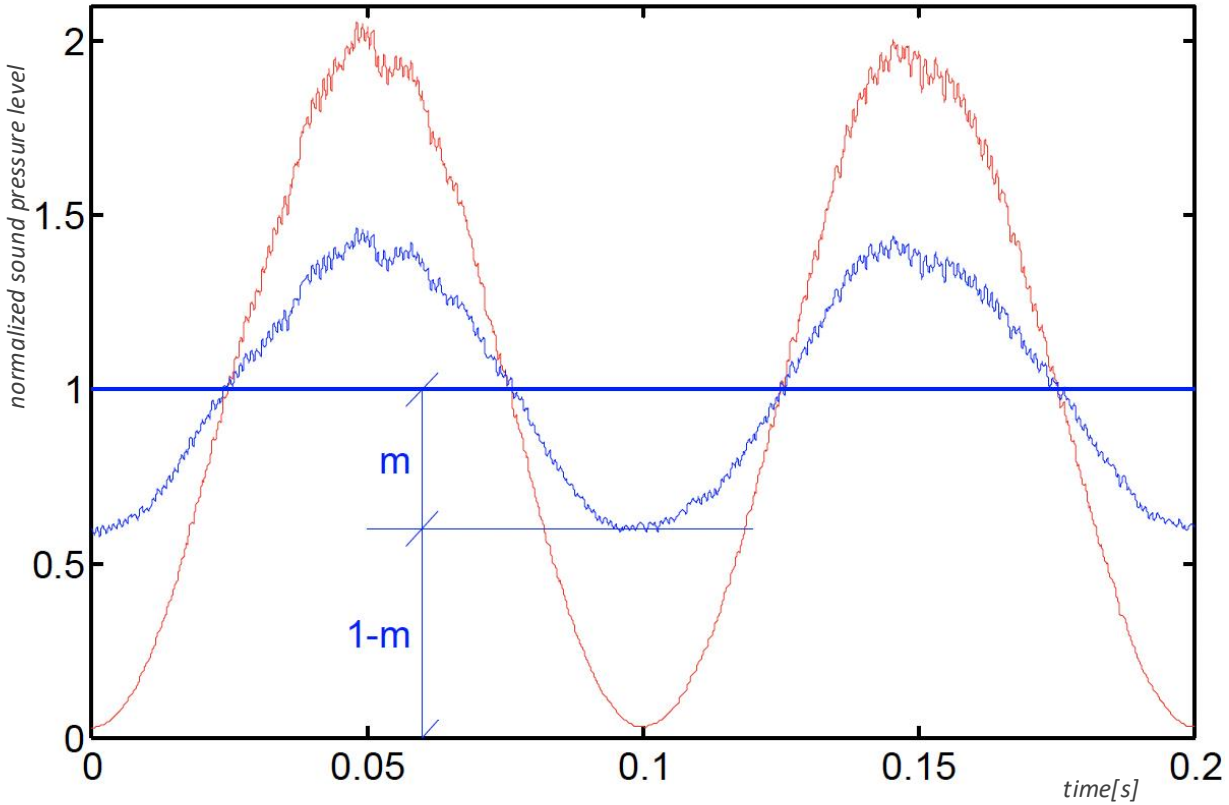




Graphics TU Delft

How to measure intelligibility ?

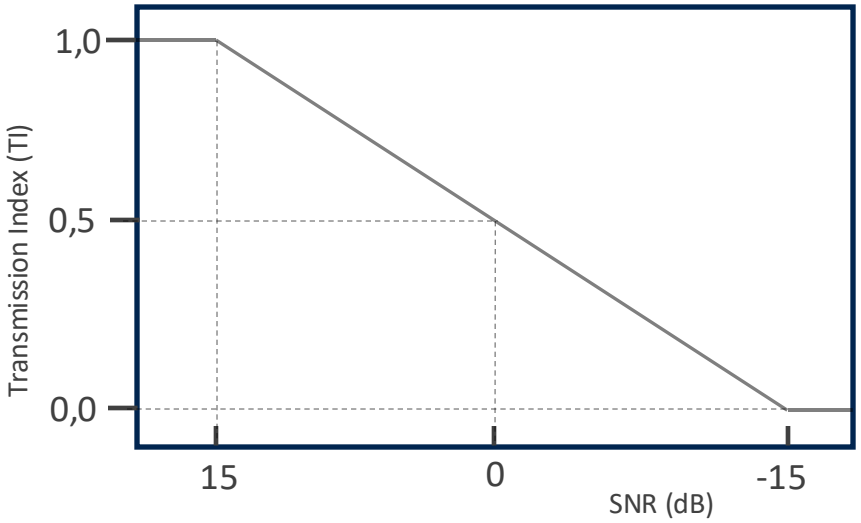
- Modulation to Transfer Index: TI



Graphics TU Delft

$$SNR = 10 \cdot \log\left(\frac{m}{1-m}\right)$$

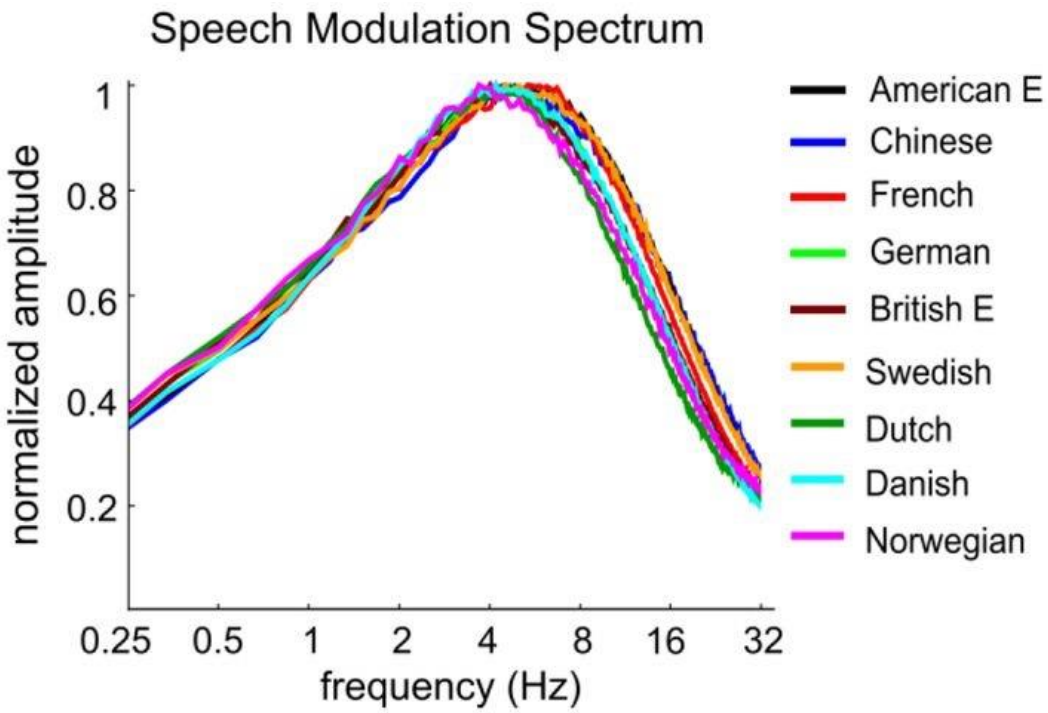
$$TI = \left(\frac{SNR + 15}{30}\right)$$

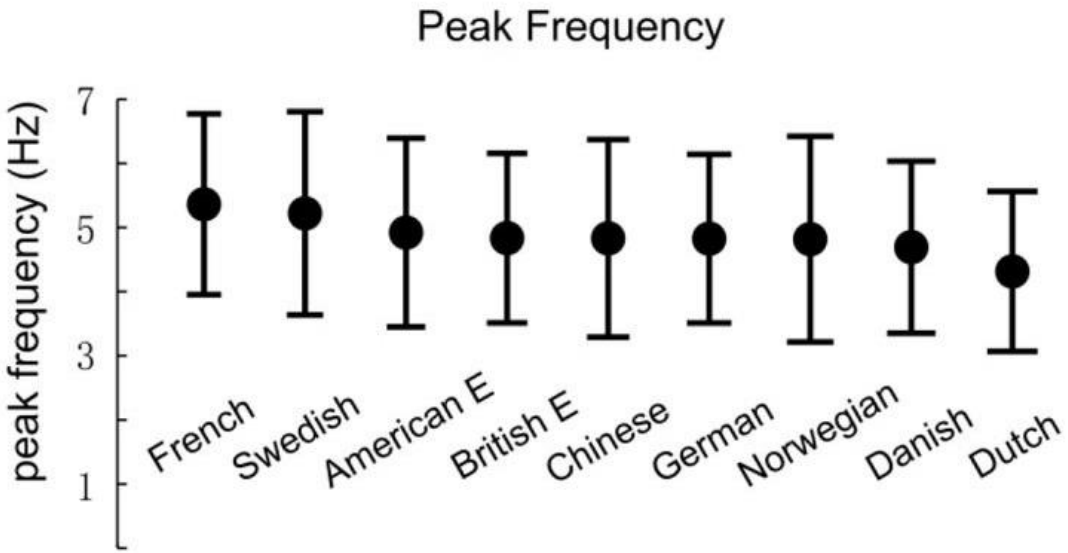


YAMAHA

How to measure intelligibility ?

- Modulation



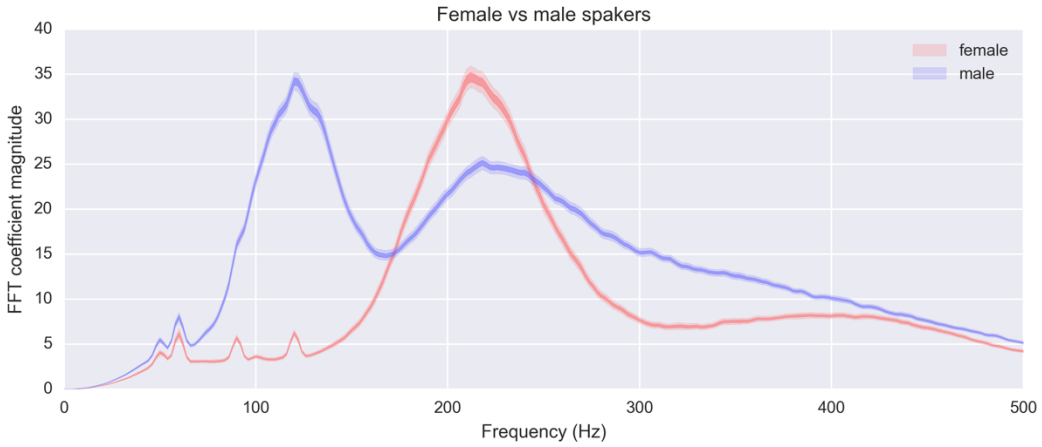


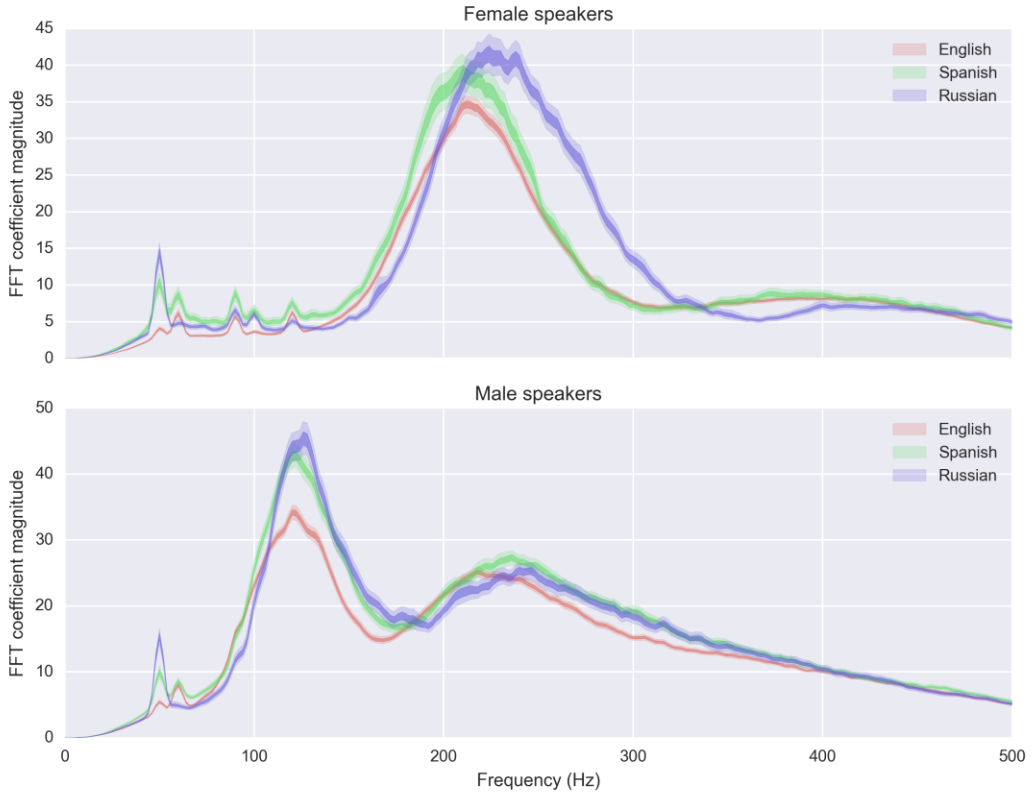
Data from Nai Ding, Aniruddh D. Patel, Lin Chen, Henry Butler, Cheng Luo, David Poeppel

YAMAHA

How to measure intelligibility ?

- Spectral content



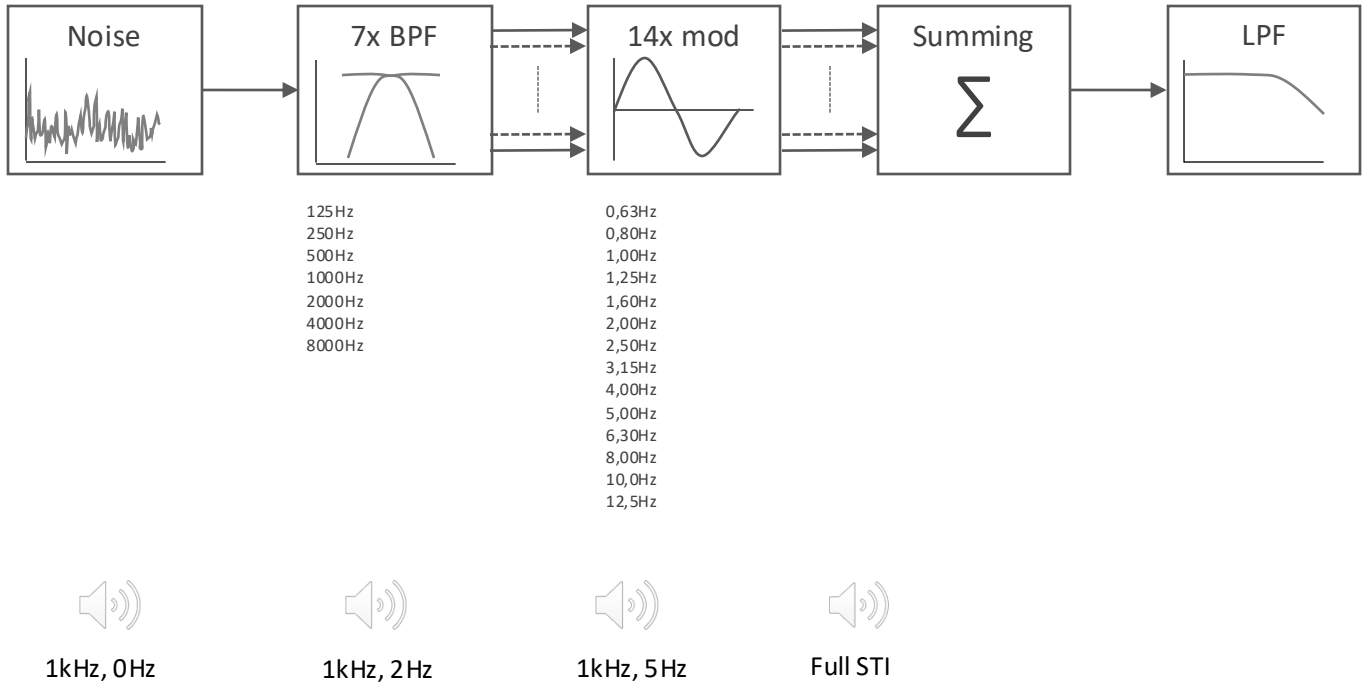


Data from Erik Bernhardsson

YAMAHA

How to measure intelligibility ?

- Joining modulation and spectral content: The STI reference signal



How to measure intelligibility ?

- The reference signal; modulation

Modulation frequency	Speech frequency (Hz) 125	250	500	1000	2000	4000	8000
0,63	v	v	v	v	v	v	v
0,8	v	v	v	v	v	v	v
1,0	v	v	v	v	v	v	v
1,25	v	v	v	v	v	v	v
1,6	v	v	v	v	v	v	v
2,0	v	v	v	v	v	v	v
2,5	v	v	v	v	v	v	v
3,15	v	v	v	v	v	v	v
4,0	v	v	v	v	v	v	v
5,0	v	v	v	v	v	v	v
6,3	v	v	v	v	v	v	v
8,0	v	v	v	v	v	v	v
10,0	v	v	v	v	v	v	v
12,5	v	v	v	v	v	v	v
Weight w_{fs}	0,129	0,143	0,114	0,114	0,186	0,171	0,143
MTI_{fs}	MTI_{125}	MTI_{250}	MTI_{500}	MTI_{1000}	MTI_{2000}	MTI_{4000}	MTI_{8000}

$$MTI_{fs} = \frac{1}{14}\sum_{0,63}^{12,5} TI_{fm}$$

$$STI = \sum_{125}^{8000}\{w_{fs} \cdot MTI_{fs}\}$$

YAMAHA

How to measure intelligibility ?

- The STI category labels

	STI
Bad	0,00 – 0,29
Poor	0,30 – 0,44
Fair	0,45 – 0,59
Good	0,60 – 0,74
Excellent	0,75 – 1,00

STI	IEC 60268-16 ed4 (2011)
0,00 - 0,36	U
0,36 – 0,40	J
0,40 – 0,44	I
0,44 - 0,48	H
0,48 – 0,52	G
0,52 – 0,56	F
0,56 – 0,60	E
0,60 – 0,64	D
0,64 – 0,68	C
0,68 – 0,72	B
0,72 – 0,76	A
0,76 – 1,00	A^+

YAMAHA

How to measure intelligibility ?

- Full STI: 14x7 measurements: 98 measurements in total (~ 15 minutes)
- STIPA, RASTI and STITEL are variations to shorten the measurement time

STIPA

Speech frequency (Hz)

Modulation frequency	125	250	500	1k	2k	4k	8k
0,71			V				
1,12	V						
1,78						V	
2,83				V			
4,53							V
6,97					V		
11,33		V					

RASTI

Speech frequency (Hz)

Modulation frequency	125	250	500	1k	2k	4k	8k
0,7					V		
1,0			V				
1,4					V		
2,0			V				
2,8					V		
4,0			V				
5,6					V		
8,0			V				
11,2					V		

OBSOLETE

STITEL

Speech frequency (Hz)

Modulation frequency	125	250	500	1k	2k	4k	8k
0,63			V				
0,8						V	
1,0		V					
1,25					V		
1,6	V						
2,0				V			
2,5							V
3,15			V				
4,0						V	
5,0		V					
6,3					V		
8,0	V						
10,0				V			
12,5							V

OBSOLETE

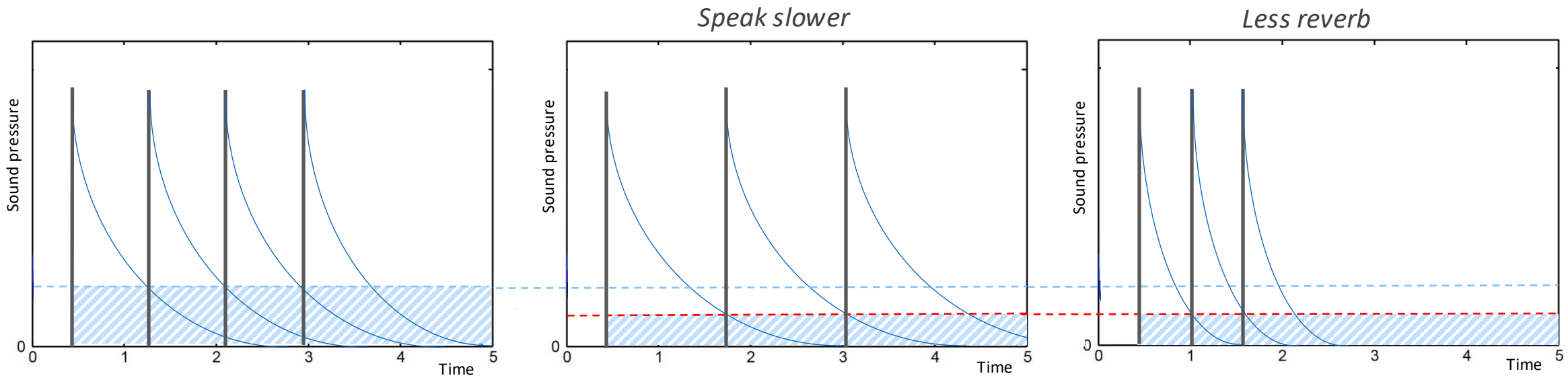
YAMAHA

When to distrust an STI?

- Big tolerances in measurements
- Strong flutter echos (especially for STIPA)
- Non-linear distortion in the system (compression, limiting, AGC)
- Strong non-flat frequency response
- Time variency (moving people, doors, street noises)

How to improve intelligibility ?

- From the math (and experience) we know reverb is most important factor
- Reverb is nothing else than noise, reducing the modulation transfer



How to improve intelligibility ?

- Two variables are important: T (Reverb time) and S/N (BGN)
- The formula for the modulation depth:

$$m(F_{mod,SN}) = \underbrace{\frac{1}{\sqrt{1+\left[\frac{2\pi F_{mod} T}{13{,}8}\right]^2}}}_{Reverb} \cdot \underbrace{\frac{1}{1+10^{-0{,}1SN}}}_{BGN}$$

YAMAHA

How to improve intelligibility ?

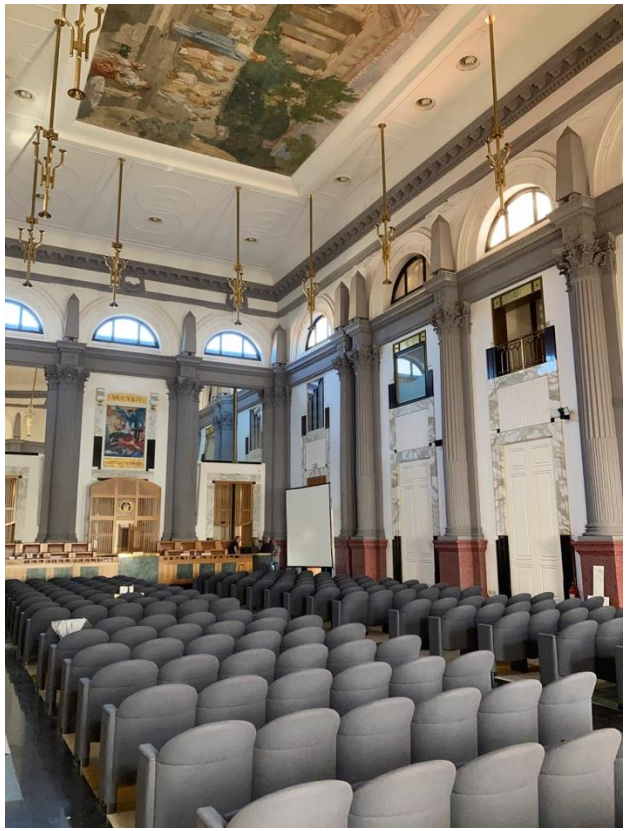
- Source
  - Voice, speaking technique, speed, articulation
  - Microphone quality, microphone directivity, microphone technique
- Acoustics
  - Avoid echoes
  - Decrease RT
- External influences
  - Background noise, streetnoises, fans
- System
  - SNR
  - Equalization
  - No distortion
  - Better speaker systems
  - Using directional speakers

Case examples

Rector Hall University Federico II, Napoli

Case examples

University Federico II Napels (IT)

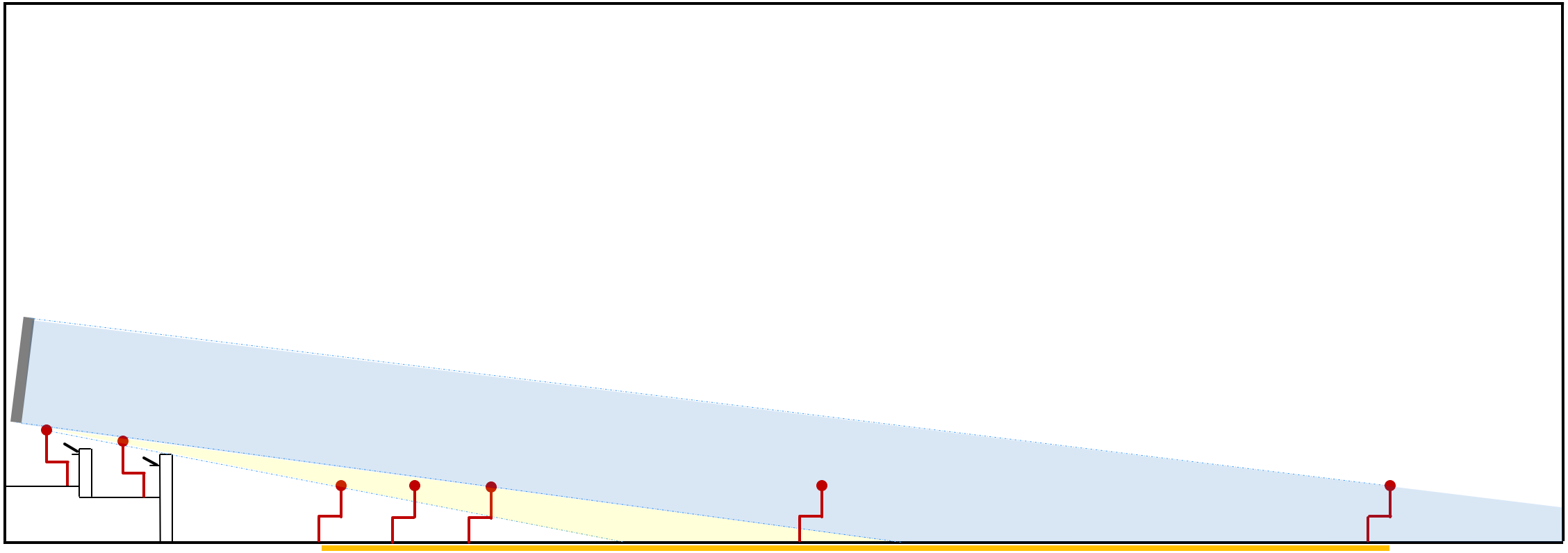


Length:	27,6 [m]
Width:	13,3 [m]
Height:	13,0 [m]
Volume:	4770 [$m^{3]}$
Material:	Marble floor Stone walls
$RT60_{[RT20,mid]}$:	3,0 [s]

YAMAHA

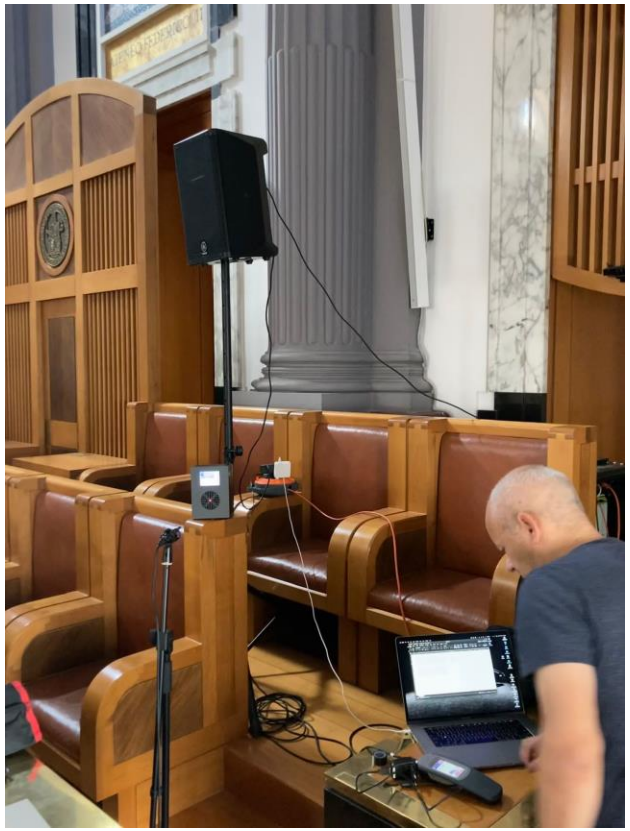
Case examples

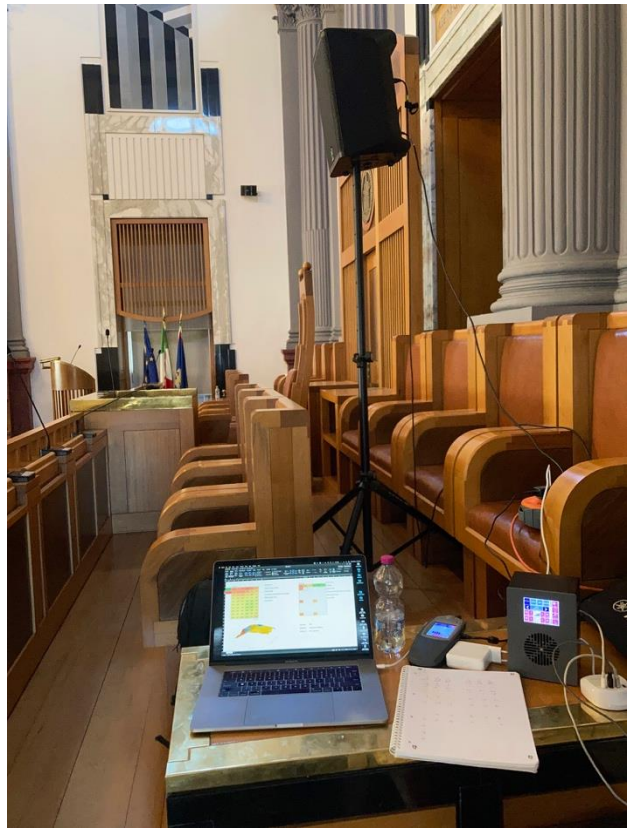
University Federico II Napels (IT)



Case examples

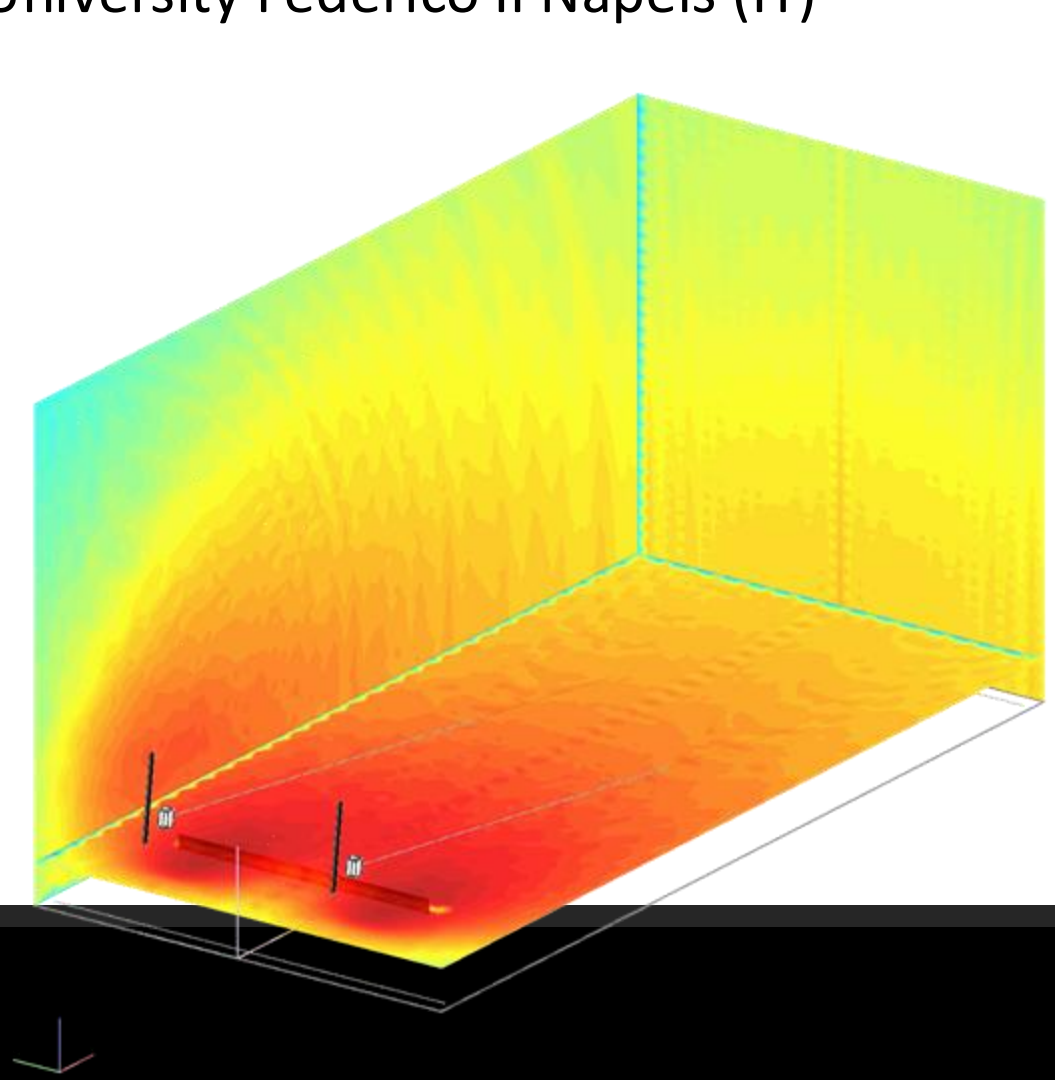
University Federico II Napels (IT)

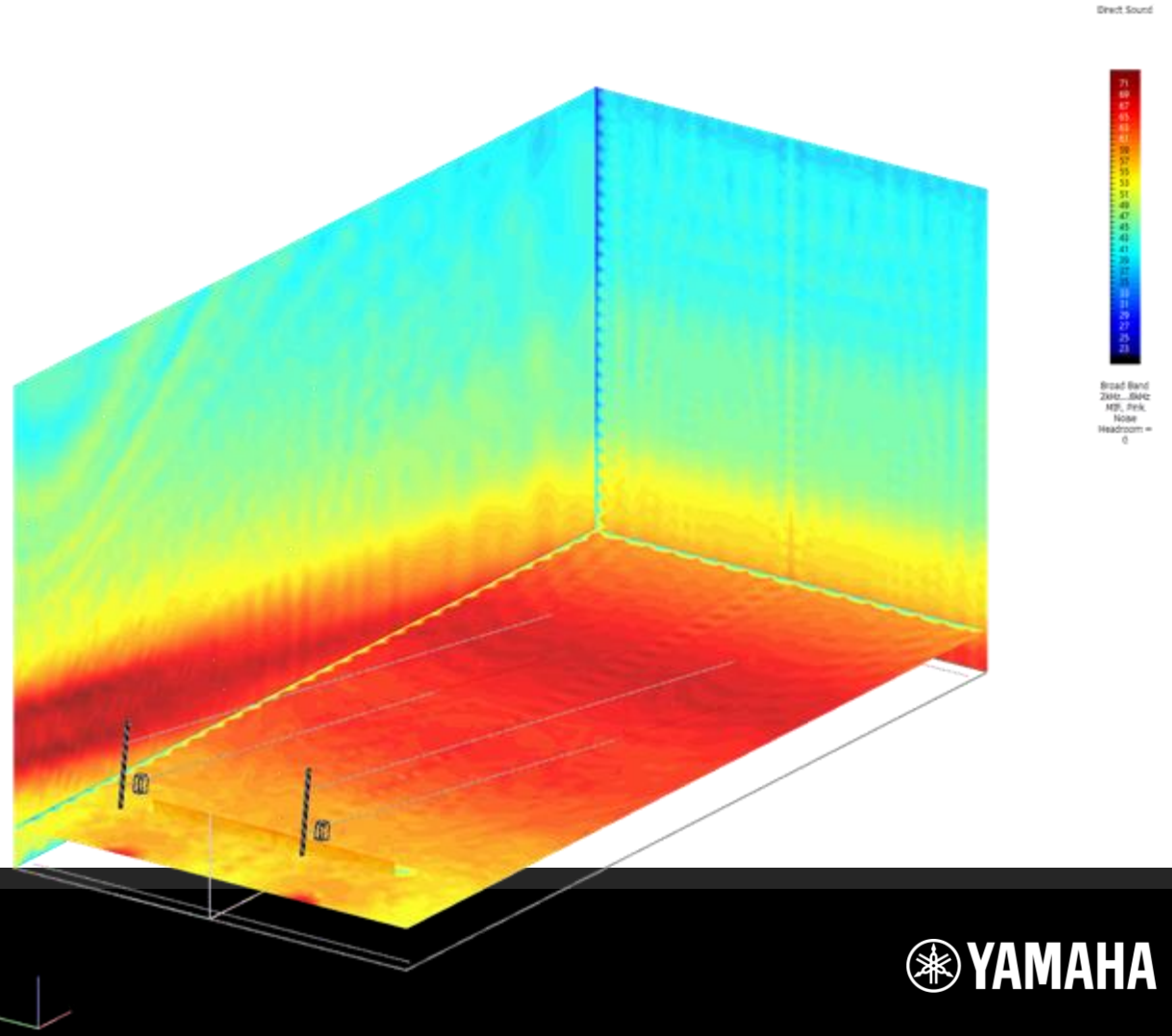




Case examples

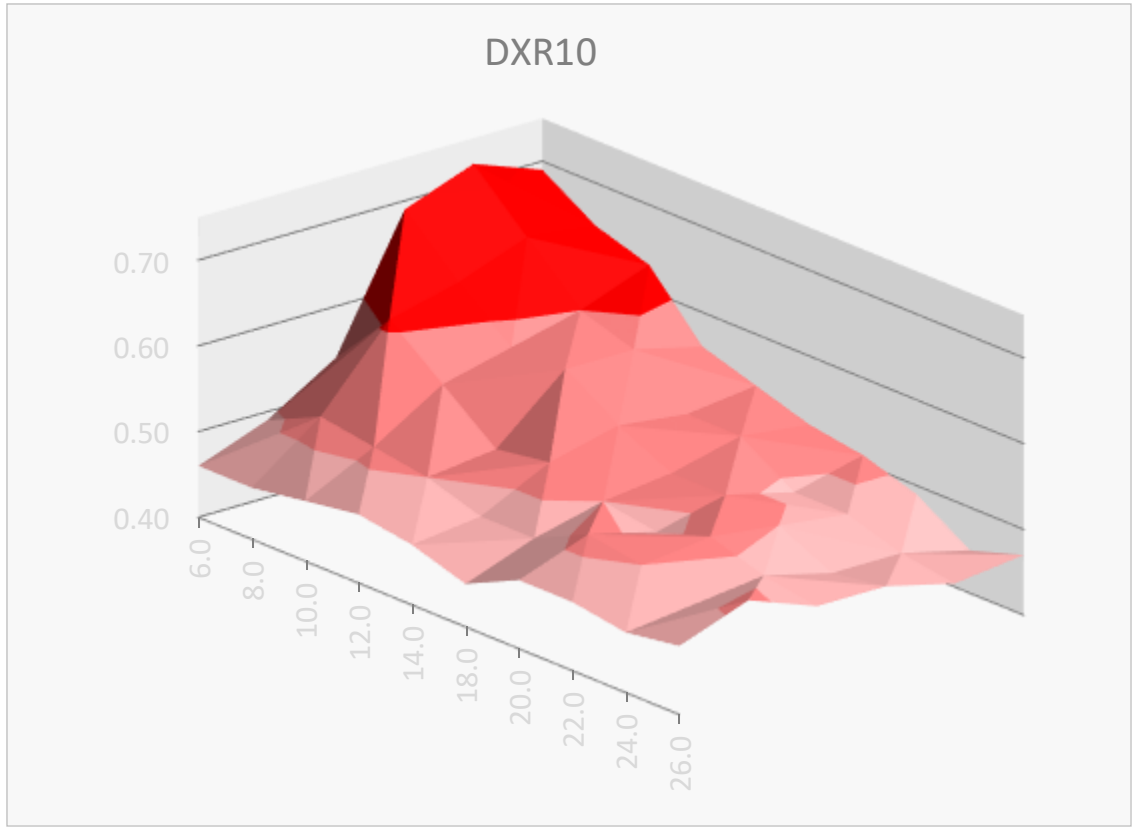
University Federico II Napels (IT)

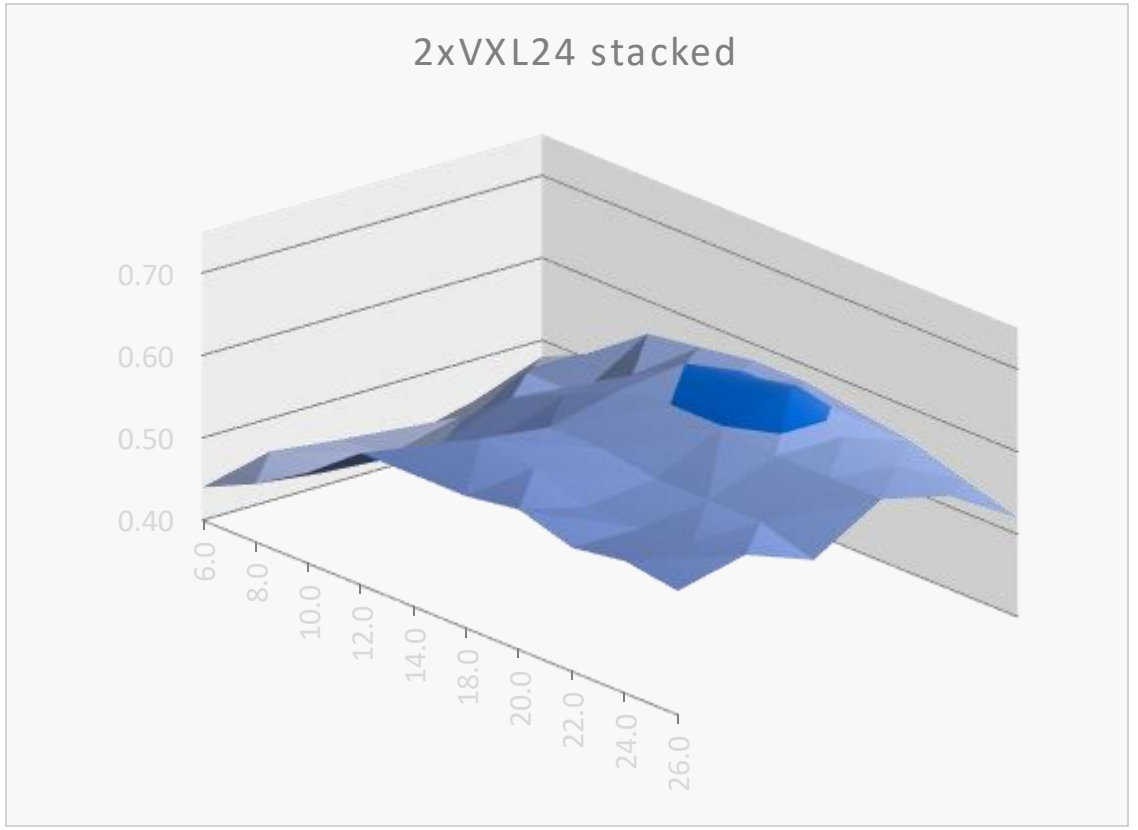




Case examples

University Federico II Napels (IT)

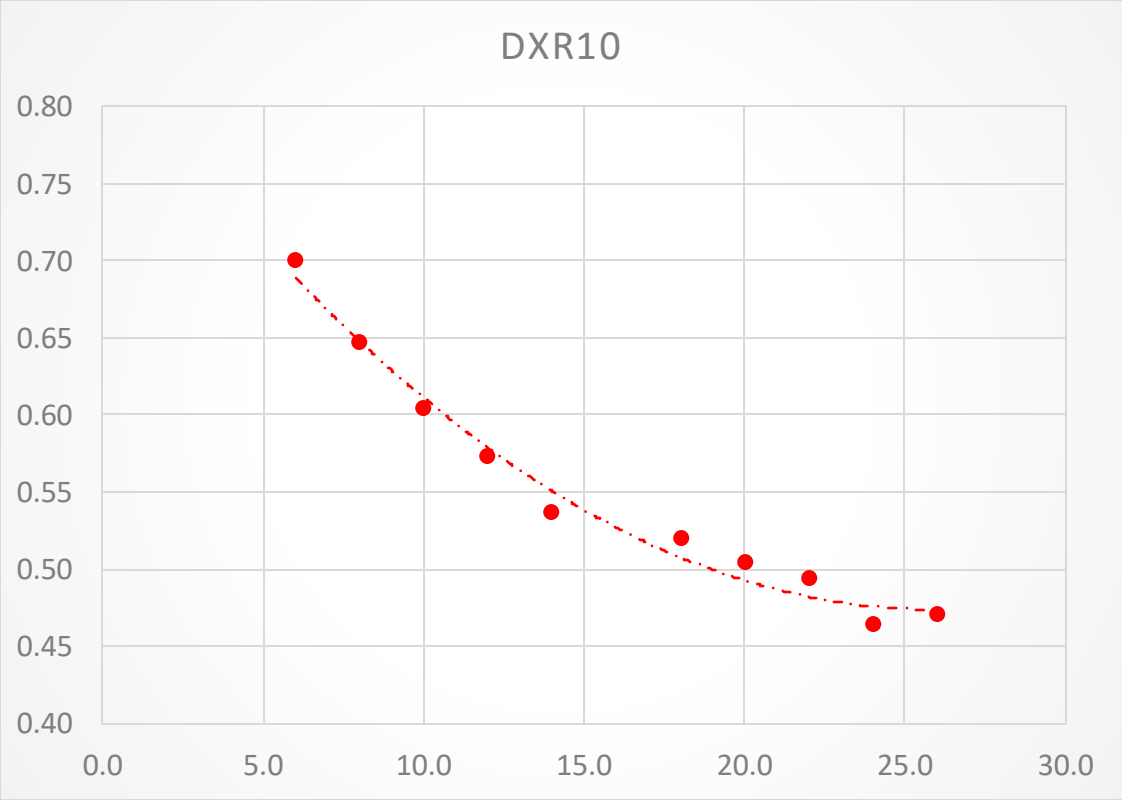


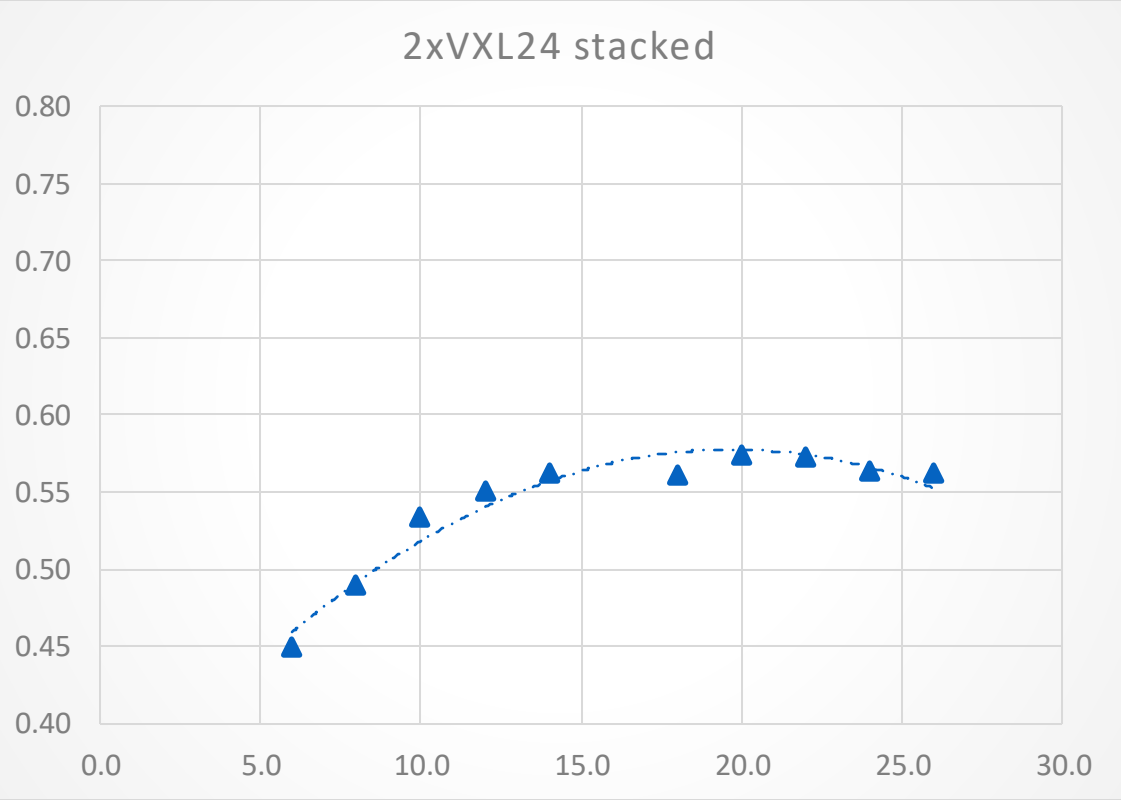


YAMAHA

Case examples

University Federico II Napels (IT)

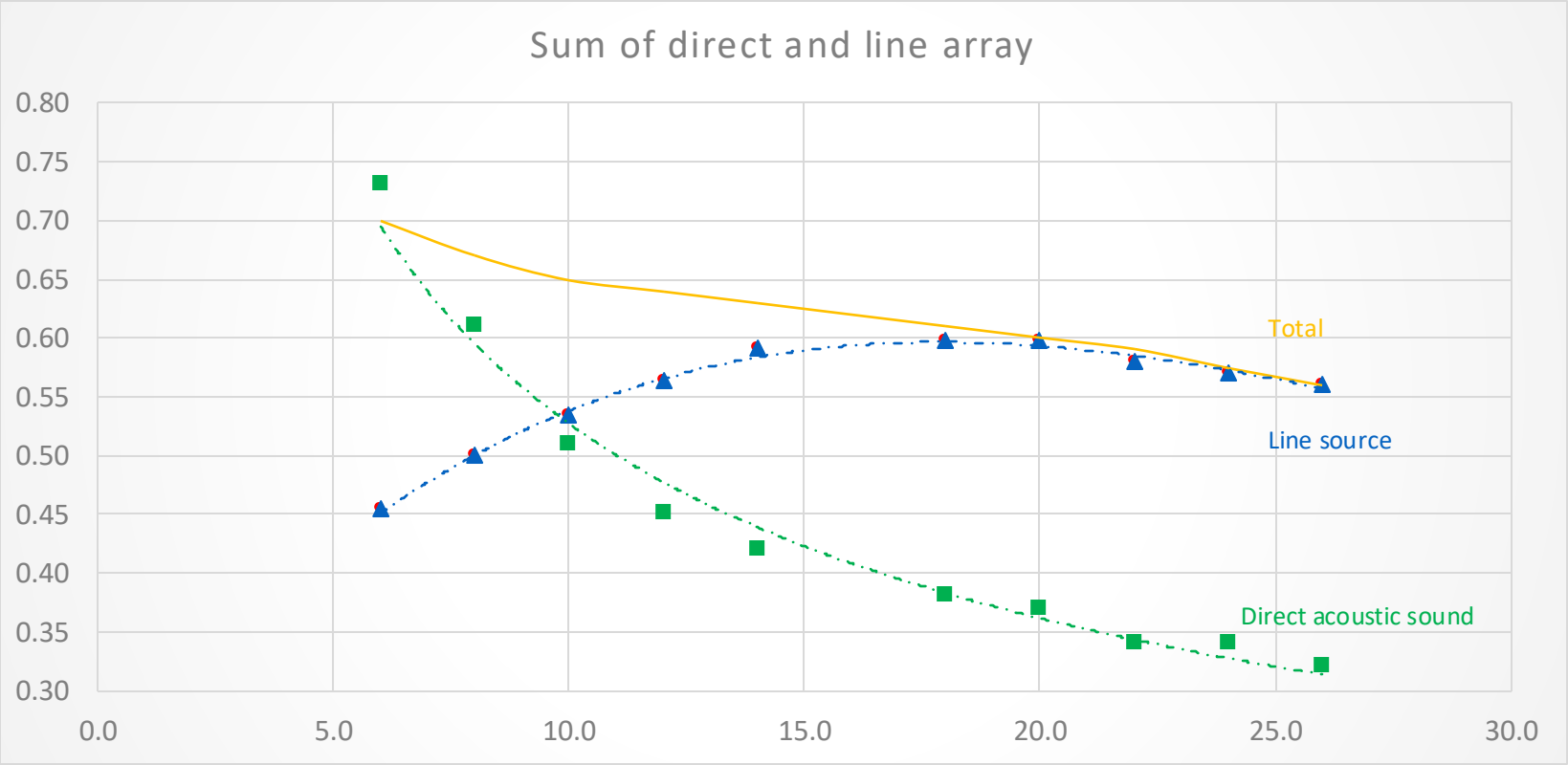




YAMAHA

Case examples

University Federico II Napels (IT)



Case examples

Steinerner Saal, Vienna

Case examples

Steinerner Saal (SDC 2025 Vienna)

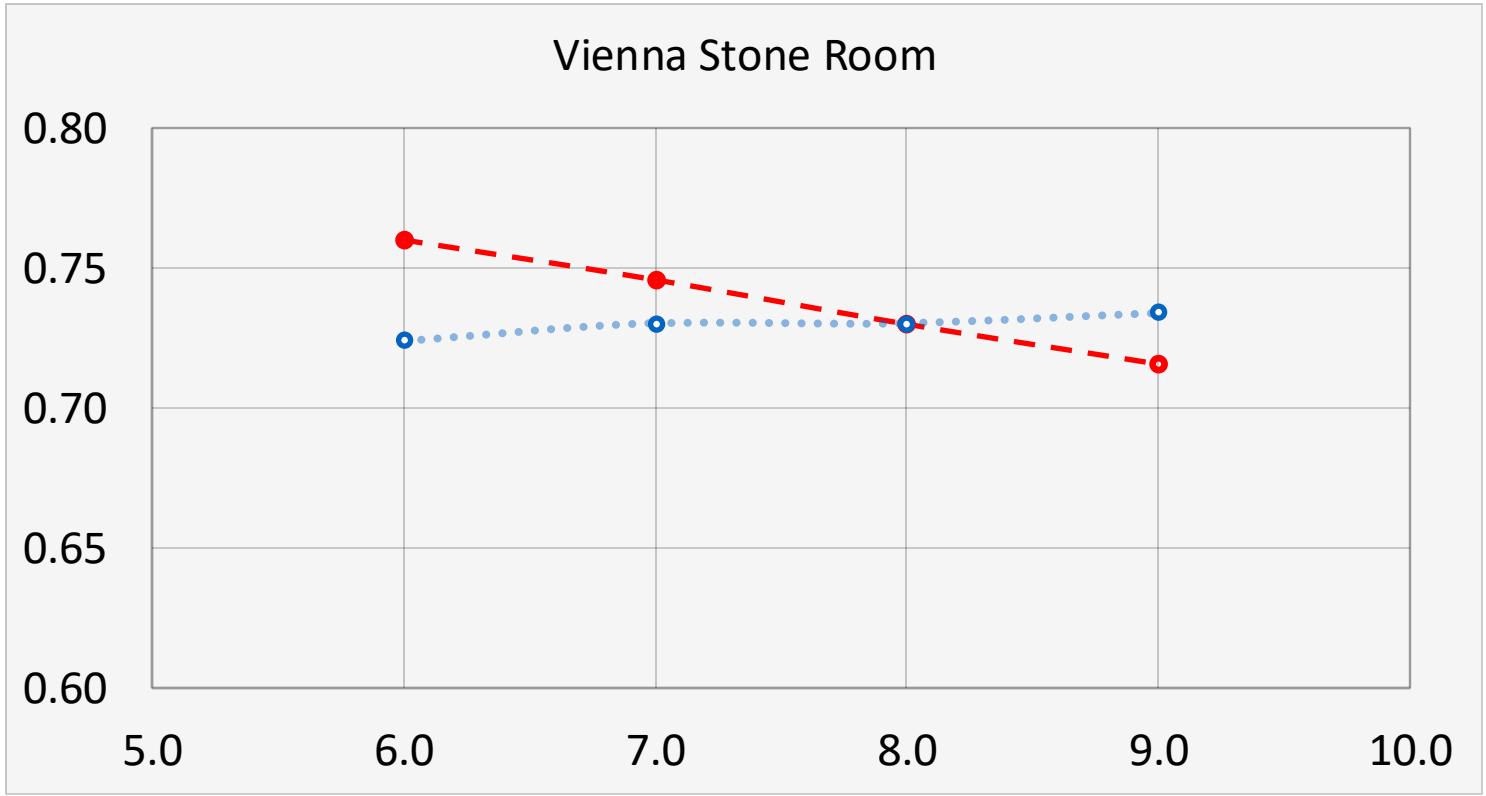
Length:	12,7 [m]
Width:	8,6 [m]
Height:	3,2 [m]
Volume:	349 [m^3]
Material:	Carpet Stone walls

$RT60_{[RT20,mid]}$: 0,8 [s]

YAMAHA

Case examples

Steinerner Saal (SDC 2025 Vienna)



YAMAHA
Make Waves

From classrooms to cathedrals

Enhancing Speech Intelligibility

Wouter Verkuijl
Yamaha Music Europe

Ruben van der Goor
Yamaha Music Europe

SDC 2025
Vienna


YAMAHA
Make Waves